

A spatial-temporal stochastic rainfall model for Auckland City: Scenarios for current and future climates

Paul Cowpertwait¹, Jim Salinger² and Brett Mullan³

¹ *Institute of Information and Mathematical Sciences, Massey University, Albany Campus, Auckland, New Zealand. Corresponding author: p.s.cowpertwait@massey.ac.nz*

² *National Institute for Water and Atmospheric Research, Newmarket, Auckland City, New Zealand.*

³ *National Institute for Water and Atmospheric Research, Greta Point, Wellington, New Zealand.*

Abstract

A spatial-temporal stochastic model of rainfall suitable for urban catchment studies is described. The model is fitted to data as part of the Auckland City Integrated Catchment Study, aimed at upgrading the Auckland City wastewater network and minimising the pollution impact on receiving watercourses. The stochastic model is based on a Neyman-Scott point process that is extended spatially using a Poisson process of rain cell origins, with cell radii following independent exponential distributions. The model is a conceptual-stochastic model, rather than a purely empirical statistical model, and incorporates features of the underlying meteorological process, albeit in an idealised way. A good fit is obtained to statistical properties sampled from historical rainfall time series at sites across Auckland City, including annual extreme values that are not used in model fitting. In addition, three future rainfall scenarios are simulated using the model, to allow for a possible change in climate due to global warming.

Introduction

The Integrated Catchment Study, a project valued at over NZ\$23 million, was a four-year study to develop a detailed understanding of the Auckland drainage system in order to minimise wastewater overflows that can cause public health and environmental problems (Sharman *et al.*, 2006). In assessing the performance of a drainage network, under a range of design options, rainfall data form an essential input to hydraulic computer models of the system (Davis *et al.*, 2007). However, rainfall data are usually limited, so in consequence engineers have sought models of rainfall to assess system performance via simulation. The traditional single-event 'design storm' approach, based on a statistical analysis of extreme values, has a number of associated problems, mainly because a single event cannot adequately represent the range of conditions that lead to high levels of pollution entering a receiving water course – see, for example, Vaes *et al.* (2001) for a discussion of design storms. The alternative to using a single event model is to use a time

series of rainfall as input to a mathematical model of the hydraulic system. This approach is computationally intensive, but with recent advances in computing technology this has become less of a concern.

There has been considerable work on developing stochastic models of rainfall suitable for urban catchment studies. For example, in the UK, the commercial software package ‘Stormpac’ was developed for this purpose (Water Research Centre, 2008), and is now widely used throughout the UK and other parts of Europe. The ‘Stormpac’ software suite, however, does not allow for spatial variation in rainfall over a catchment. To allow for this, a full spatial-temporal stochastic model was developed and can be used to simulate statistically representative rainfall series at multiple sites (e.g., Cowpertwait, 1995). Simulations from this model have been used to address various urban pollution management problems in industry, for example, the Thames Tideway Tunnels project (Cowpertwait, 2006). However, only the Auckland catchment study used simulations of plausible future rainfall scenarios that allow for a change in climate.

The objective of the present paper is to describe the spatial-temporal model that was used for Auckland City, which has been summarised in a number of conference proceedings that place the model in the context of the wider catchment study (e.g., see Davis *et al.*, 2007; Lockie *et al.* 2006). In this paper we assess the model fit to statistical properties of the historical rainfall data, and provide information on how the future climate scenarios for Auckland were generated.

Assembly of data

Rainfall data (1992-2002) at the five-minute level of aggregation for twelve sites in Auckland City catchment were provided by Metrowater (Table 1). Data sampled at one-hour time steps were downloaded from the NIWA climate database (Penny, 1996). The locations of these sites were Whenuapai, Albert Park, and Auckland Airport. In addition, a long record of daily data for Albert Park was downloaded to assess possible long-term trends in the mean rainfall and proportion of dry days.

Table 1 – Rainfall records used to fit the stochastic model

Site Name	Latitude (°)	Longitude (°)	Scale (h)	Period (years)
Abatoir	-36.924	174.849	1/12	1992-2002
Albert Park (1)	-36.850	174.767	1/12	1992-2002
Alexandra Park	-36.892	174.778	1/12	1992-2002
Avondale	-36.899	174.697	1/12	1992-2002
Churchill Park	-36.855	174.875	1/12	1992-2002
Cox's Bay	-36.849	174.719	1/12	1992-2002
Mt Albert	-36.884	174.715	1/12	1992-2002
Mt Roskill	-36.905	174.742	1/12	1992-2002
Okahu Bay	-36.853	174.814	1/12	1992-2002
Onehunga	-36.923	174.785	1/12	1992-2002
Tamaki	-36.890	174.857	1/12	1992-2002
Bassant	-36.908	174.809	1/12	1992-2002
Albert Park (2)	-36.850	174.767	1	1966-1985
Auckland Airport	-37.005	174.783	1	1966-1985
Whenuapai	-36.780	174.636	1	1966-1985
Albert Park (3)	-36.850	174.767	24	1869-2000

To reduce the sampling error of the parameter estimates, the model can be fitted to pooled data, provided the pooled data all fall in the same homogenous region (e.g., see Institute of Hydrology, 1999). Standard t-tests that compared the mean rainfall for each calendar month showed that monthly mean values for Auckland Airport differed significantly (with p-values under 5%) from those for Auckland City (Albert Park), using data from the common period of 1966-1985. The means for each calendar month in the Whenuapai series, however, were not significantly different from those obtained for the Auckland City series. Hence, the hourly data from Albert Park were pooled with the data from Whenuapai. In addition, the five-minute series from the Auckland catchment were aggregated to the hourly level and pooled with the data from Albert Park and Whenuapai. This resulted in a pooled series of hourly data based on historical data for the periods 1966-1985 and 1992-2002. The pooled series covered both positive and negative phase shifts of the Interdecadal Pacific Oscillation index – a climate index that has been proposed for capturing long-term climate variation and persistence (e.g., see Zhang *et al.*, 1997). The pooled data were therefore suitable for estimating the parameters of the stochastic model for the current climate.

The long record of daily data from Albert Park (Table 1) was used to assess long-term trends in the mean and in the proportion of dry days, to determine whether these trends should be extrapolated to 2050. Polynomial functions of time (in years) were fitted by least-squares regression to the annual means and annual proportion of dry days. There was marginal statistical evidence of trends (with p-values for coefficients under 10%) in both series, although in both cases a further ARIMA analysis indicated that the trends could be stochastic, i.e. accounted for by high serial correlation. Hence, extrapolation

of the polynomial trends (or ARIMA model predictions) to 2050 would be unreliable. A similar analysis indicated that the Interdecadal Pacific Oscillation Index had stochastic trends which could not be reliably extrapolated into the future. Hence, for future climate predictions, extrapolations based on General Circulation Models (GCMs) of the atmosphere were preferred, essentially because scenarios were required for a considerable time step ahead into the future.

Definition of a spatial-temporal stochastic model

Let the arrival times $\{T_i\}$ of storm origins occur in a Poisson process with rate λ . A storm consists of a stochastic process of rain cells, denoted by the set of random variables $\{(U_{ij}, V_{ij}), S_{ij}, L_{ij}, X_{ij}, R_{ij}\}$, where for the i -th storm:

1. $\{(U_{ij}, V_{ij})\}$ forms a two-dimensional Poisson process with rate ζ ;
2. (U_{ij}, V_{ij}) and R_{ij} form discs in two-dimensional space, where (U_{ij}, V_{ij}) is the disc centre and R_{ij} is the disc radius, which is taken to be an independent exponential random variable with parameter ϕ ;
3. S_{ij} is the arrival time of the j -th cell in the i -th storm, where $S_{ij} - T_i$ are independent exponential random variables with parameter β , so that the cell arrival times $\{S_{ij}\}$ form a Neyman-Scott point process;
4. L_{ij} is the lifetime of the j -th cell, which is taken to be an independent exponential random variable with parameter η , so that the j -th cell in the i -th storm terminates at a random time $S_{ij} + L_{ij}$;
5. X_{ij} is a random variable representing the intensity of the j -th cell in the i -th storm, where X_{ij} remains constant throughout the cell lifetime and over the area of the disc $\{(U_{ij}, V_{ij}), R_{ij}\}$.

6. The total intensity at time t and location $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, $Y(\mathbf{x}, t)$ is the sum of the intensities of all cells alive at time t and overlapping location \mathbf{x} .

For the purpose of model fitting and simulation, the cell intensity X is taken to be a Weibull (α, θ) random variable with r -th moment given by: $E(X^r) = \theta^r \Gamma(1+r/\alpha)$. Furthermore, for each storm, the number of cells C that overlap a point in two-dimensional space \mathbb{R}^2 is a Poisson random variable with mean $\mu_C = 2\pi\xi/\phi^2$. Thus, the model is summarised by the following set of parameters:

- λ^{-1} – the mean time (h) between adjacent storm origins;
- β^{-1} – the mean waiting time (h) for a cell origin after a storm origin;
- μ_C – the mean number of rain cells per storm overlapping a point in the catchment;
- η^{-1} – the mean cell lifetime (h);
- α – shape parameter for cell intensity;
- θ – scale parameter (mm h^{-1}) for cell intensity;
- ϕ^{-1} – the mean cell radius (km).

The stochastic model is a conceptual model that has rain cells represented by rectangular pulses in time and circular discs in space. Whilst these assumptions are unrealistic at small scales, upon aggregation the model can provide realistic profiles and time series properties. Time series properties, up to third-order, are given in Cowpertwait (1995, 1998). Derived properties include the mean, autocovariance, crosscorrelation, and third moment, which are functions of the above model parameters and the sampling interval (or aggregation level) h . For example, when $h = 24$ (hours), the daily mean, daily autocovariance and daily third moment are available as functions of the model parameters.

Fitted model

Following Cowpertwait *et al.* (2002), the model is fitted to the pooled hourly series for each calendar month by minimising the following sum of squares:

$$\sum_{\hat{f} \in S} \left\{ \left(1 - f / \hat{f} \right)^2 + \left(1 - \hat{f} / f \right)^2 \right\}$$

where S is the set of dimensionless time-series properties: coefficient of variation, autocorrelation at lag 1, and skewness, each evaluated at both the hourly and daily levels of aggregation. The sample estimates \hat{f} of these properties are found for each calendar month by pooling the available data for each month over all years in the record of hourly data. The parameters are estimated for each month separately and stationarity is assumed over a monthly period. Since a minimization procedure is used, an exact fit to sample properties in the summation is not expected – the use of the reciprocal term in the summation ensures the fitted values are not biased above or below the sample estimates. The procedure provides estimates of all model parameters except the scale parameter θ , since dimensionless functions are used which do not depend on θ . The scale parameter θ is estimated after the other parameters directly from the hourly sample mean $\hat{\mu}$, and is given by: $\hat{\theta} = \hat{\mu} \hat{\eta} / \left\{ \hat{\lambda} \hat{\mu}_C \Gamma(1+1/\hat{\alpha}) \right\}$ (see Cowpertwait *et al.* 2002). Since the model is fitted separately to each month, this procedure results in 12 estimates for each model parameter (Table 2), which ensures that the model properties approximately match the observed mean, variance, lag 1 autocorrelation, and sample skewness at both the hourly and daily levels for all calendar months. Finally, for each month, the cell radius parameter (ϕ) is estimated by minimising the difference between the sample

Table 2 – Stochastic model parameter estimates for Auckland City

Month	$\hat{\lambda}$ (h ⁻¹)	$\hat{\mu}_C$ per storm	$\hat{\beta}$ (h ⁻¹)	$\hat{\eta}$ (h ⁻¹)	$\hat{\alpha}$	$\hat{\theta}$ (mm h ⁻¹)	$\hat{\phi}$ (km ⁻¹)
1	0.00578	14.3	0.0738	2.159	0.594	1.496	0.140
2	0.00390	16.0	0.0659	0.603	0.487	0.446	0.098
3	0.00530	18.3	0.1048	1.393	0.549	0.911	0.219
4	0.00866	11.4	0.1169	1.117	0.688	1.294	0.089
5	0.01000	26.9	0.1257	0.969	0.392	0.155	0.191
6	0.01193	42.1	0.1054	0.862	0.364	0.073	0.109
7	0.01069	27.8	0.1026	1.17	0.550	0.479	0.147
8	0.02345	7.83	0.2312	0.754	0.547	0.395	0.111
9	0.01323	16.1	0.1464	1.141	0.541	0.453	0.124
10	0.00999	18.4	0.1187	1.333	0.533	0.484	0.109
11	0.00891	20.5	0.1216	1.336	0.519	0.520	0.125
12	0.00718	35.8	0.1377	2.077	0.436	0.363	0.227

hourly crosscorrelation and the equivalent model function, using data from all available sites in the catchment. Note that the scale parameter can be modified to account for any non-homogeneity in the rainfall process or adjusted to allow for a change in the mean rain cell intensity, which might happen as a result of climate change.

The fitted parameters are given in Table 2, where a seasonal variation in the estimates can be seen. For example, over the winter months λ increases and θ and η both decrease, corresponding to an increase in low-intensity large-scale frontal weather. Conversely, in the summer months, the cells tend to be shorter and of higher intensity (η and θ both increase), which corresponds to an increase in convective rainfall.

Overall, the model has a close fit to the sample values (Figs. 1-6), although the hourly autocorrelations tend to be slightly over-estimated by the model (Fig. 3). Also, there is a slight under-estimation of the daily skewness (Fig. 6), but as peak flows are more

likely to be dependent on hourly depths, the slight lack of fit in the distribution tail of the daily rainfall may not be of practical importance. In general, these discrepancies could be assessed for their practical significance using hydraulic flow simulation models.

The sample and fitted cross-correlations are shown for January and July, from which a satisfactory fit is evident, particularly as one parameter (ϕ) is used to fit the sample cross-correlations at all sites (Figs. 7 and 8). Discrepancies evident in the July series are probably due to missing pairs of values in the calculation of hourly cross-correlation. The January fitted values may decay slightly too slowly, possibly due to the assumption of spatial homogeneity across the catchment. For large regions, the model can be re-parameterised to account for limited storm sizes (e.g., see Leonard *et al.*, 2008), although for the Auckland City catchment this would not be of much practical value since the estimated mean cell radius is of similar scale to the catchment size.

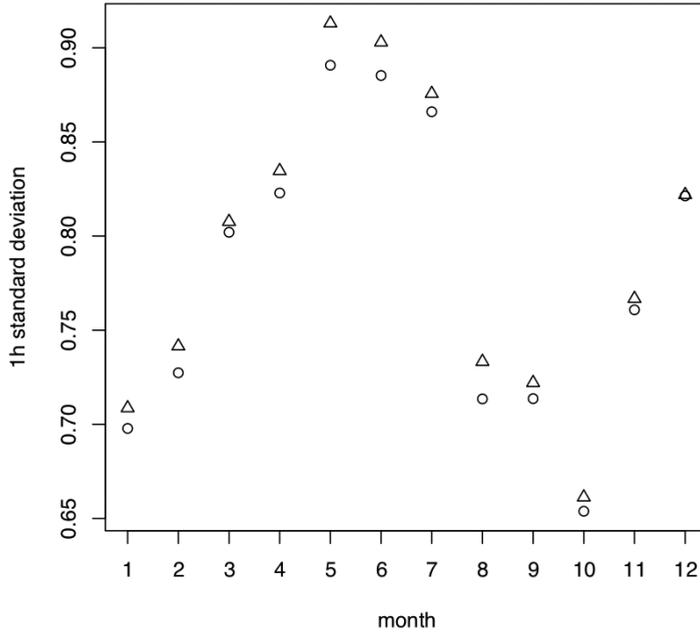


Figure 1 – Standard deviation of hourly series: ○ Observed; △ Fitted.

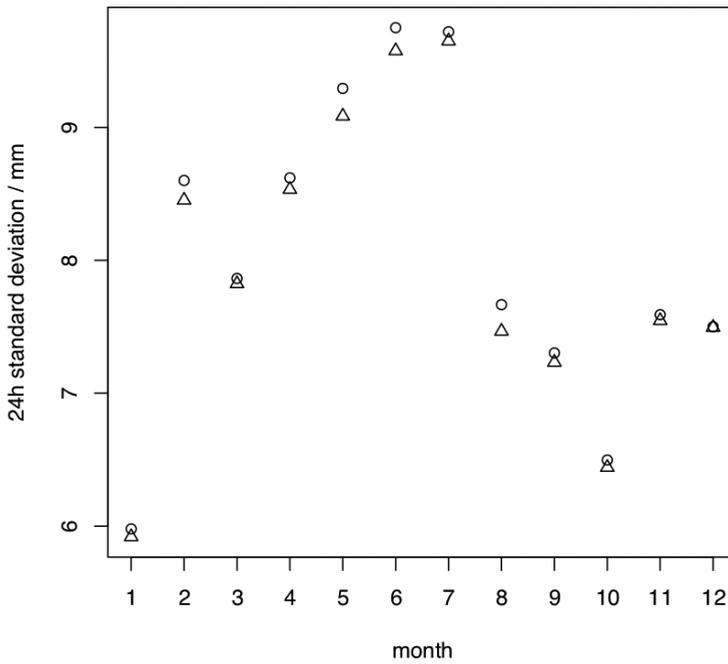


Figure 2 – Standard deviation of daily series: ○ Observed; △ Fitted.

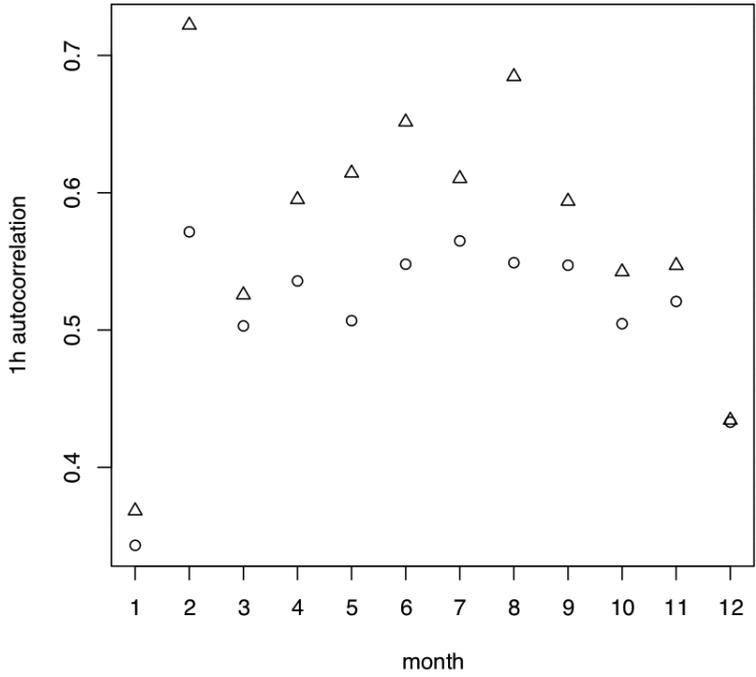


Figure 3 – Autocorrelation (lag 1) of hourly series: ○ Observed; △ Fitted.

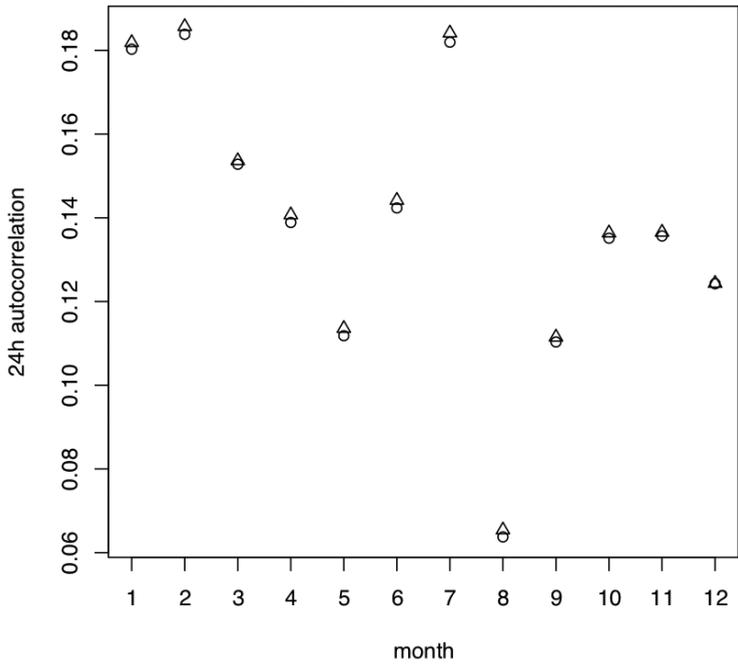


Figure 4 – Autocorrelation (lag 1) of daily series: ○ Observed; △ Fitted.

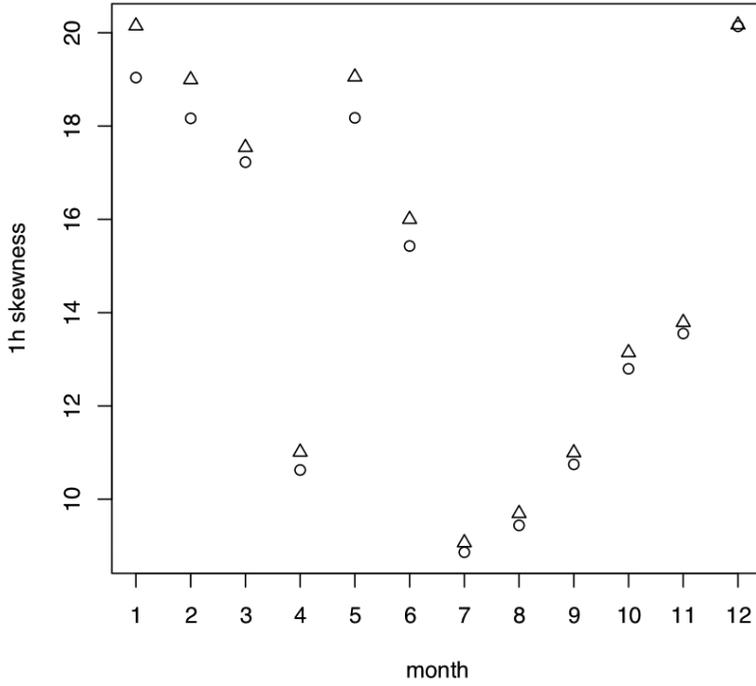


Figure 5 – Skewness of hourly series: ○ Observed; △ Fitted.

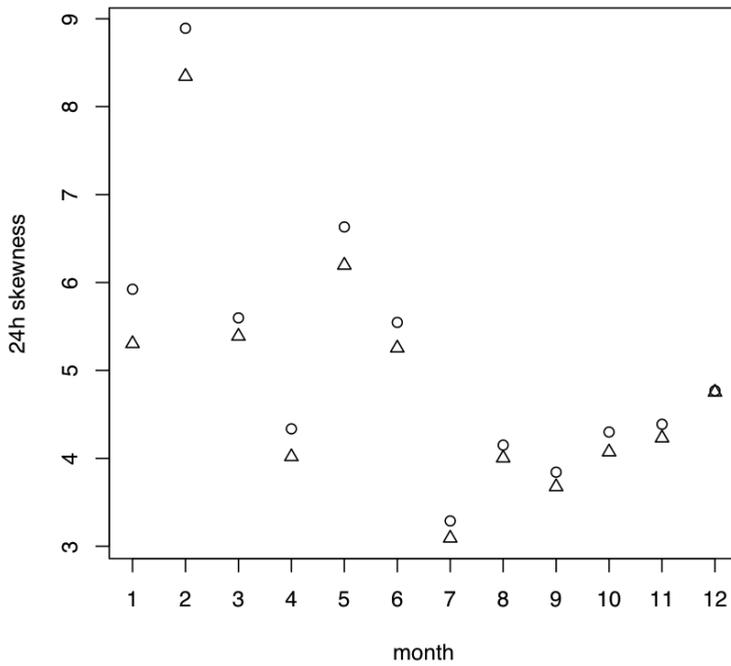


Figure 6 – Skewness of daily series: ○ Observed; △ Fitted.

Cross-correlations for January

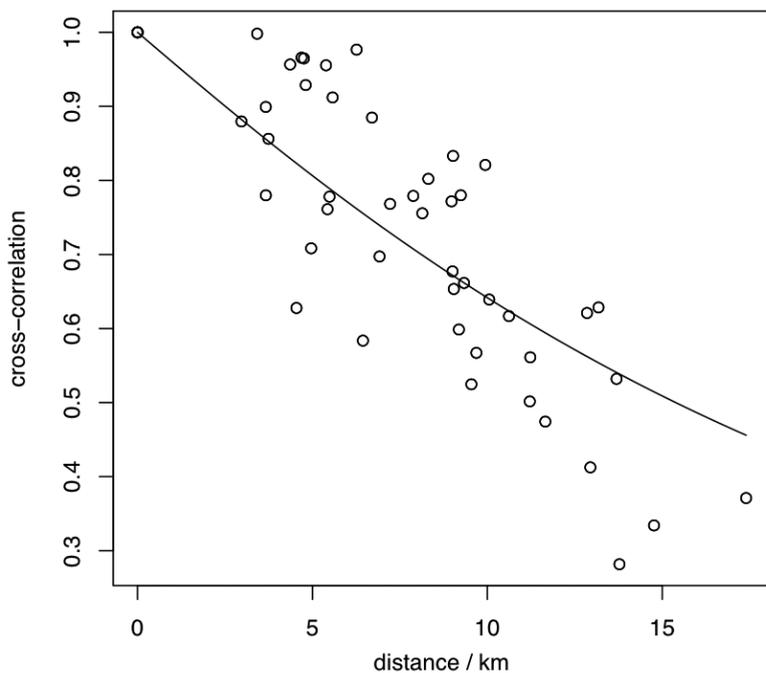


Figure 7 – Observed and fitted hourly cross-correlations against distance for January

Cross-correlations for July

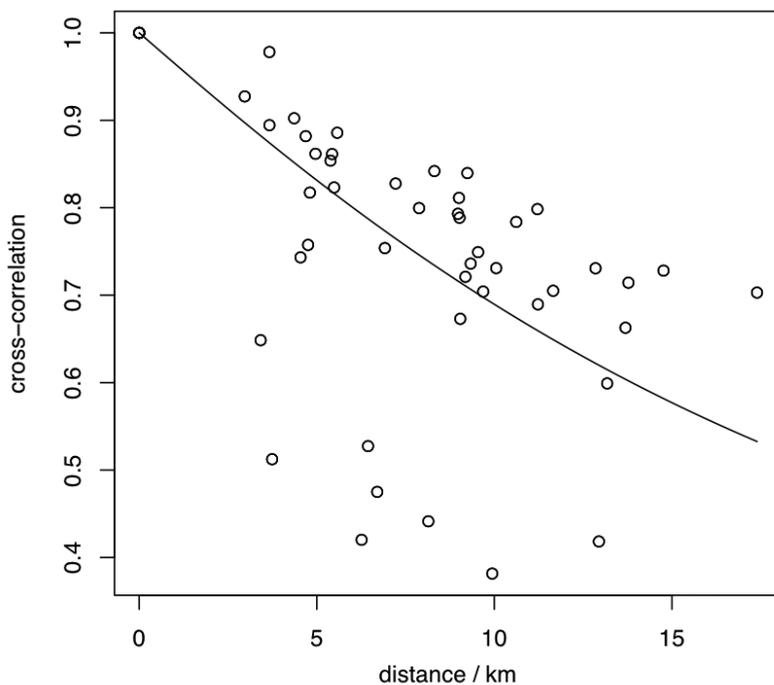


Figure 8 – Observed and fitted hourly cross-correlations against distance for July.

Model validation

In model validation it is appropriate to use properties that were not used in the fitting procedure. Annual maxima were selected, as these were not used to fit the model and are of importance in the intended application. The annual maxima were extracted at both the hourly and daily levels from the Albert Park (1966-85) and the five-minute series (1992-2002; Table 1). The median values (across sites) of the annual maxima were used for the five-minute series, as some of these data had large numbers of missing values, which would bias the estimates of annual maxima. This resulted in 30 annual maxima, which were ordered and plotted against the standardised Gumbel variate (Fig. 9).

One hundred years of hourly data were simulated for the twelve sites across Auckland

City (Table 1), using the fitted spatial-temporal stochastic model. For each of the 12 sites, the simulated series were divided into three periods of 30 years (to give equal number to the observed), and for each 30-year period the annual maxima were extracted and ordered. This resulted in 3×12 values at each Gumbel ordinate. The simulated maxima are shown with the historical maxima in Figure 9. This procedure was repeated at the daily level of aggregation and the equivalent results are given in Figure 10. In general, a good fit is obtained at the hourly and daily levels of aggregation, giving confidence in the fitted model and simulation algorithm (Figs. 9 and 10). At the daily level of aggregation, there is a slight tendency to underestimate extreme values in the lower range (Fig. 10), which may be due to the slight

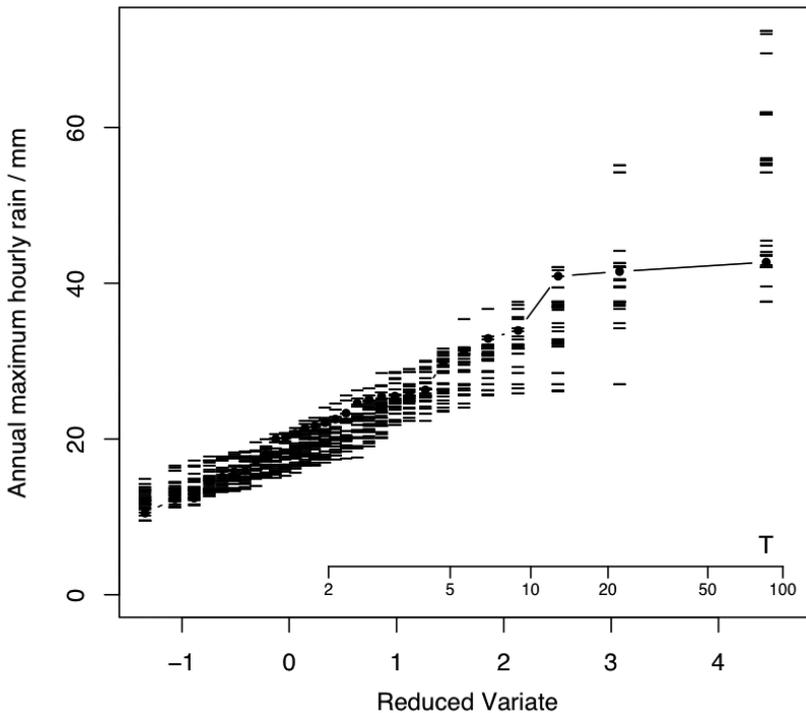


Figure 9 – Ordered annual hourly maximums for pooled historical (• and line) and simulated (–) data.

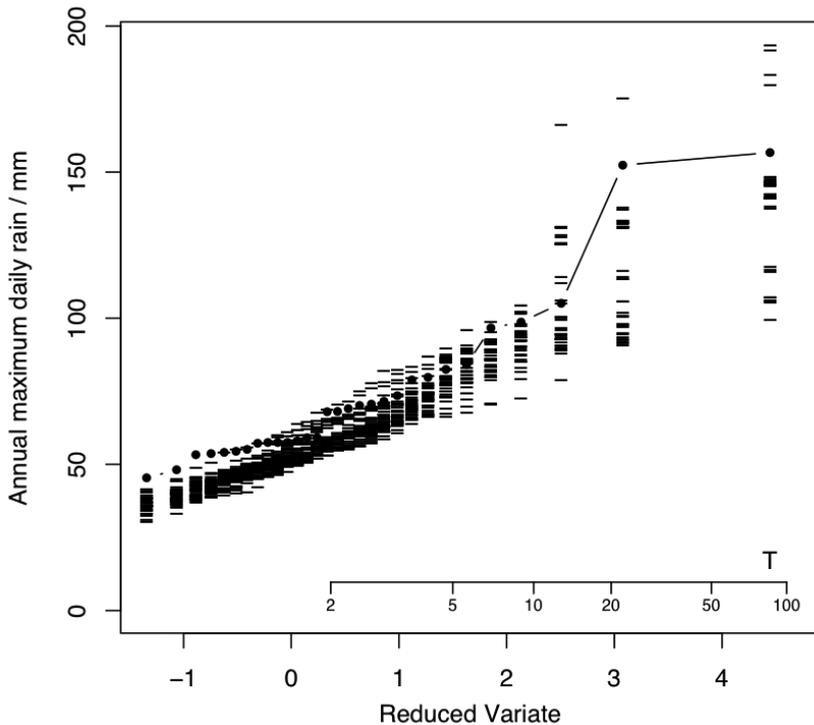


Figure 10 – Ordered annual daily maximums for pooled historical (• and line) and simulated (–) data.

under-estimation of the daily skewness (Fig. 6). Overall, however, the plots indicate a good fit has been obtained, especially as the extreme values tend to be sensitive to small departures in goodness-of-fit.

Future climate scenarios

Mullan *et al.* (2001) assessed a range of simulations from General Circulation Models for a future climate in New Zealand. The average of four GCM model simulations predicts an increase in mean rainfall of about 5%, which is likely to affect rainfall intensity (Salinger *et al.*, 2001; Mullan and Salinger, 2002). Based on Mullan and Salinger (2002), the parameter estimates for three future rainfall scenarios, all resulting in a 5% increase in the mean rainfall, are found using the functions for the stochastic model:

Future scenario 1.

For each month, the hourly mean rainfall μ_1 is increased by 5%, leaving the wet and dry periods unchanged. This is achieved by increasing the estimated scale parameter by 5%: $\theta' = 1.05\hat{\theta}$.

Future scenario 2.

For each month, the proportion of dry days is increased by 5%, using a modified estimate of the storm rate $\hat{\lambda}$, obtained from the relationship, $\lambda' = \hat{\lambda} \log(1.05)/I(24)$ where $I(24)$ is an integral expression used to calculate the proportion of dry days (Cowpertwait, 1995). The hourly mean rainfall μ_1 is then increased by 5%, using the original estimates with the revised estimate λ' in the equation that estimates the scale parameter: $\theta' = 1.05\mu_1\hat{\eta}/\{\lambda'\hat{\mu}_c\Gamma(1+1/\hat{\alpha})\}$.

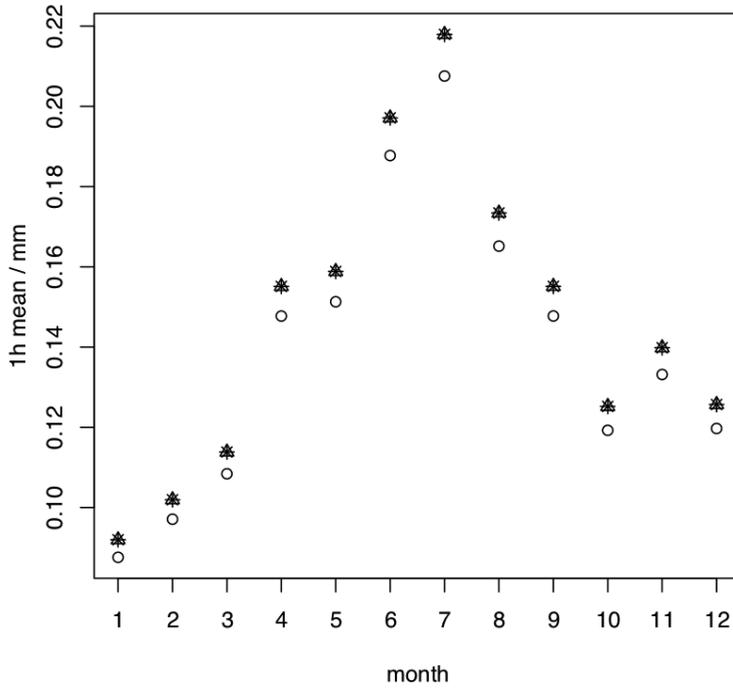


Figure 11 – Fitted hourly mean rainfall (○ Current; △ Future scenario 1; + Future scenario 2; × Future scenario 3).

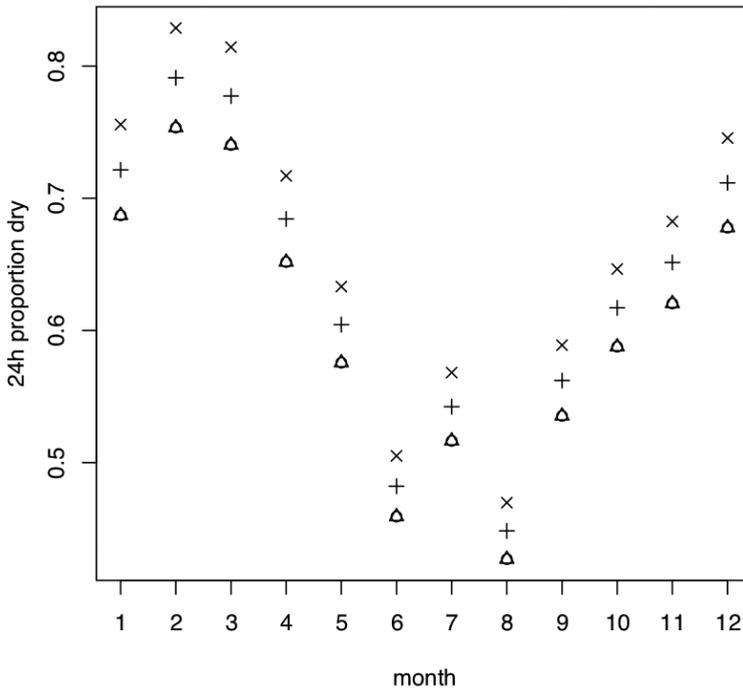


Figure 12 – Fitted proportion of dry days (○ Current; △ Future scenario 1; + Future scenario 2; × Future scenario 3).

Future scenario 3.

A procedure similar to 2 above is used for each month, but with the proportion of dry days increased by 10% using $\lambda' = \hat{\lambda} \log(1.1) / I(2.4)$. The hourly mean rainfall μ_1 is then increased by 5% using the original estimates with this revised estimate $\hat{\lambda}$ in $\theta' = 1.05\mu_1\hat{\eta} / \{\lambda'\hat{\mu}_C\Gamma(1+1/\hat{\alpha})\}$.

The resulting changes to the mean and proportion of dry days are shown in Figures 11 and 12. For each of the above scenarios, one hundred years of data were generated using the stochastic model. Since historical data does not yet exist for 2050, clearly it is not possible to validate the future scenario simulations. However, the results were compared with rainfall simulations from other climate studies for Auckland. Annual maxima for the current and the future climates were extracted from the 100 years of simulated data, ordered and plotted against the standardised Gumbel variate (Figs. 13 and 14). These compared favourably with those predicted by Salinger *et al.* (2001), and hence the simulated future scenarios were retained. As anticipated by Mullan and Salinger (2002), future scenario 3 produced the largest number of extreme values. Thus, a conservative hydraulic design could be based on this scenario.

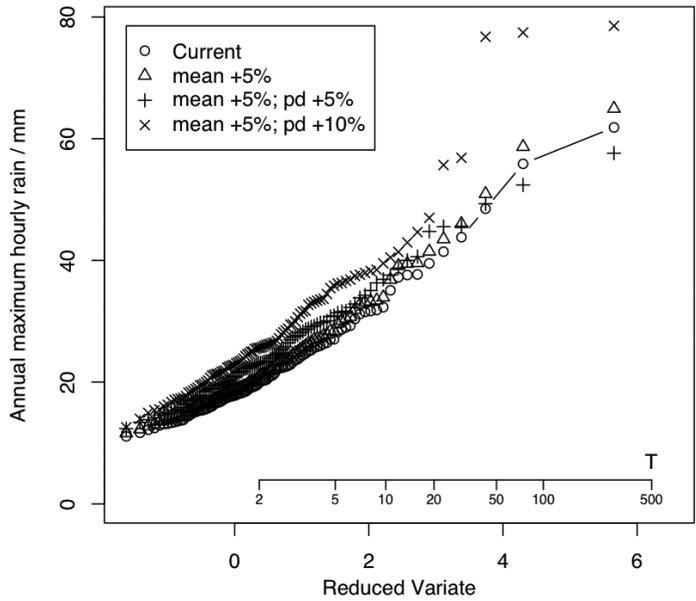


Figure 13 – Simulated ordered annual hourly maximum rainfall for current and future climate scenarios.

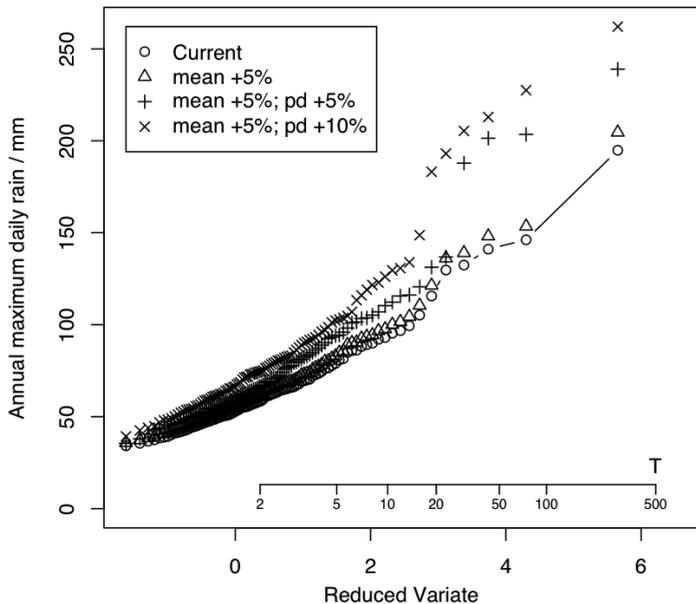


Figure 14 – Simulated ordered annual daily maximum rainfall for current and future climate scenarios.

Conclusions

Overall, the stochastic spatial-temporal model was found to fit closely a range of observed statistical properties. These include properties that were not used in the fitting procedure, and which are important in the intended application, which thus provides some confidence in the fitted model. Hence, subject to satisfactory hydraulic flow simulation tests, the model is recommended for use in urban drainage studies across Auckland City.

Acknowledgements

The authors thank Auckland City Council (ACC) and Metrowater for permission to present the work in this paper. Discussion and interpretations are those of the authors and should not be construed as those of ACC or Metrowater. Comments from the referees helped improve the paper and are gratefully acknowledged.

References

- Akaike, H. 1974: A new look at statistical model identification. *IEEE Transactions on Automatic Control* 19: 716-722.
- Chiswell, S.; Bowen, M.; Mullan, B. 2001: Climate change: New Zealand in a warming world. *Water and Atmosphere* 9: 10-11.
- Cowpertwait, P. 1995: A generalized spatial-temporal model of rainfall based on a clustered point process. *Proceedings of the Royal Society of London, Series A*, 450: 163-175.
- Cowpertwait, P. 1998: A Poisson-cluster model of rainfall: High order moments and extreme values. *Proceedings of the Royal Society of London, Series A*, 454: 885-898.
- Cowpertwait, P. 2006: A spatial-temporal point process model for the Thames catchment, London. *Journal of Hydrology* 330: 586-595.
- Cowpertwait, P.; Kilsby, C.; O'Connell, P. 2002: A space-time Neyman-Scott model of rainfall: Empirical analysis of extremes. *Water Resources Research* 38(8): 1-14.
- Cowpertwait, P.; Lockie, T.; Davies, M. 2006: A stochastic spatial-temporal disaggregation model of rainfall. *New Zealand Journal of Hydrology* 45(1): 1-12.
- Davis, M.; Kinley, P.; Salinger, M.; Mullan, A.; Cowpertwait, P.; James, M.; Sharman, B.; Paterson, G. 2007: Rainfall input for master urban drainage planning: The integrated catchment study of Auckland City, New Zealand. In Proceedings of the World Environmental and Water Resources Congress 2007 Tampa, Florida, USA. American Society of Civil Engineers.
- Institute of Hydrology 1999: *Flood estimation handbook*. Wallingford. UK.
- Leonard, M.; Lambert, M.; Metcalfe, A.; Cowpertwait, P. 2008: A space-time Neyman-Scott rainfall model with defined storm extent. *Water Resources Research* 44: W09402, doi:10.1029/2007WR006110.
- Lockie, T.; Cowpertwait, P.; Davis, M.; Kinley, P.; Paterson, G. 2006: Evaluation of historic and stochastic rainfall in network modelling in the integrated catchment study of Auckland, New Zealand. In World Water and Environmental Resources Congress 2003, Philadelphia, Pennsylvania, USA. American Society of Civil Engineers.
- Mullan, A.; Salinger, J. 2002: Scenarios of future daily rainfall for Auckland for stochastic modelling. Report to Metrowater Ltd, Auckland.
- Mullan, A.; Wratt, D.; Renwick, J. 2001: Transient model scenarios of climate change for New Zealand. *Weather and Climate* 21: 3-34.
- Penny, A. 1996: *Climate data base user's manual*. Wellington, NIWA.
- Salinger, J.; Mullan, A.; Thompson, C.; McKerchar, A.; Griffiths, G.; Timperley, M. 2001: Climate Analysis for Auckland City: Factors Affecting Water Supply and Discharge to 2050. NIWA Report AK01019 to Metrowater Ltd, Auckland.
- Sharman, B.; Paterson, G.; Brown, R.; Davis, M.; Kinley, P.; Ranchhod, S. 2006: The Integrated Catchment Study of Auckland City (New Zealand): Overview. In Proceedings of the World Environmental and Water Resources Congress 2006, Omaha, Nebraska, USA. American Society of Civil Engineers.

- Vaes, G.; Willems, P.; Barlamont, J. 2001: Rainfall input requirements for hydrological calculations. *Urban Water* 3: 107-112.
- Water Research Centre 2008: Stormpac user guide – version 4. WRc plc: Swindon, UK.
- Zhang, Y.; Wallace, J.; Battisti, D. 1997: Enso-like interdecadal variability: 1900-93. *Journal of Climate* 10: 1004-1020.